Samantha Gregoryk
CPTS 315
November 23rd, 2020
Ananth Jillepalli

EA1

Bagging (weka.classifiers.meta.Bagging). You will use decision tree (weka.classifiers.trees.j48) as the base supervised learner. Try trees of different depth (1, 2, 3, 5, 10) and different sizes of bag or ensemble, i.e., number of trees (10, 20, 40, 60, 80, 100). Compute the training accuracy and testing accuracy for different combinations of tree depth and number of trees; and plot them. List your observations.

Voting

Depth	Trees	Training Accuracy	Testing Accuracy
1	10	0.967	0.920
2	10	0.967	0.920
3	10	0.964	0.920
5	10	0.964	0.920
10	10	0.964	0.920

Depth	Trees	Training Accuracy	Testing Accuracy
1	20	0.967	0.920
2	20	0.967	0.920
3	20	0.961	0.920
5	20	0.961	0.920
10	20	0.961	0.920

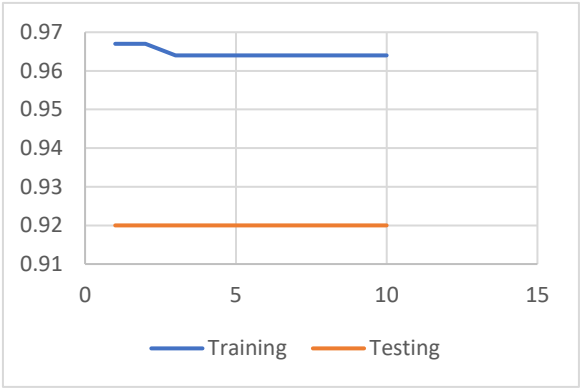
Depth	Trees	Training Accuracy	Testing Accuracy
1	40	0.967	0.920
2	40	0.967	0.920
3	40	0.967	0.910
5	40	0.967	0.900
10	40	0.967	0.900

Depth	Trees	Training Accuracy	Testing Accuracy
1	60	0.967	0.920
2	60	0.967	0.920
3	60	0.967	0.920
5	60	0.967	0.920
10	60	0.967	0.920

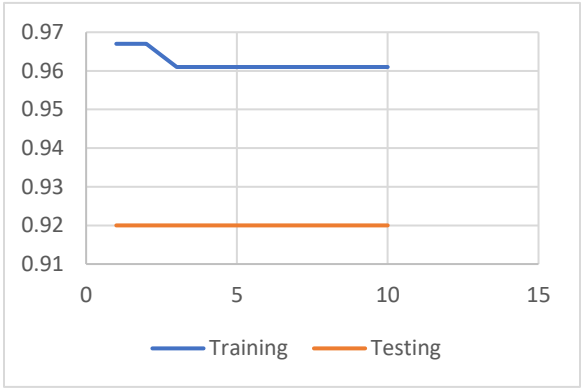
Depth	Trees	Training Accuracy	Testing Accuracy
1	80	0.967	0.920
2	80	0.967	0.920
3	80	0.964	0.900
5	80	0.964	0.900
10	80	0.964	0.900

Depth	Trees	Training Accuracy	Testing Accuracy
1	100	0.967	0.920
2	100	0.967	0.920
3	100	0.967	0.920
5	100	0.967	0.910
10	100	0.967	0.910

10 Trees



20 Trees



40 Trees

60 Trees

80 Trees

100 Trees

Ionosphere

Depth	Trees	Training Accuracy	Testing Accuracy
1	10	0.779	1.000
2	10	0.886	1.000
3	10	0.886	1.000
5	10	0.886	1.000
10	10	0.882	1.000

Depth	Trees	Training Accuracy	Testing Accuracy
1	20	0.776	1.000
2	20	0.882	1.000
3	20	0.886	1.000
5	20	0.875	1.000
10	20	0.875	1.000

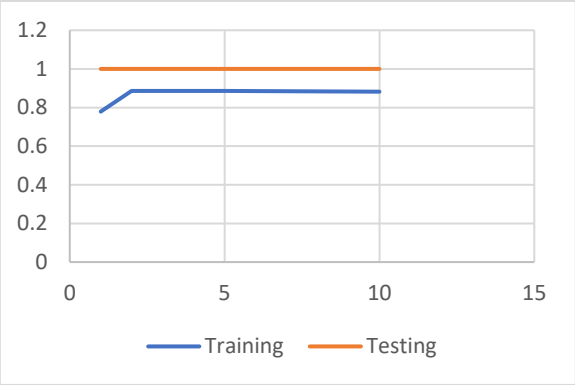
Depth	Trees	Training Accuracy	Testing Accuracy
1	40	0.768	1.000
2	40	0.886	1.000
3	40	0.886	1.000
5	40	0.878	1.000
10	40	0.878	1.000

Depth	Trees	Training Accuracy	Testing Accuracy
1	60	0.772	1.000
2	60	0.878	1.000
3	60	0.878	1.000
5	60	0.875	1.000
10	60	0.875	1.000

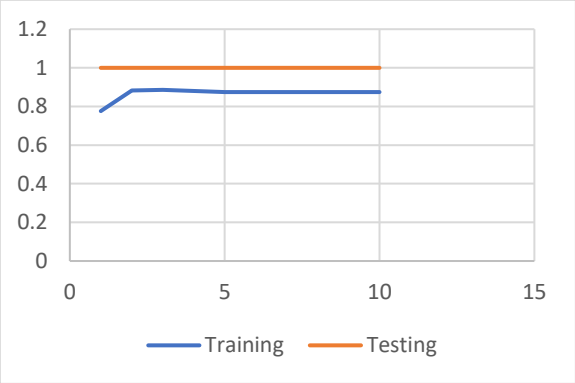
Depth	Trees	Training Accuracy	Testing Accuracy
1	80	0.783	1.000
2	80	0.875	1.000
3	80	0.878	1.000
5	80	0.875	1.000
10	80	0.875	1.000

Depth	Trees	Training Accuracy	Testing Accuracy
1	100	0.779	NA
2	100	0.878	NA
3	100	0.878	NA
5	100	0.859	NA
10	100	0.867	NA

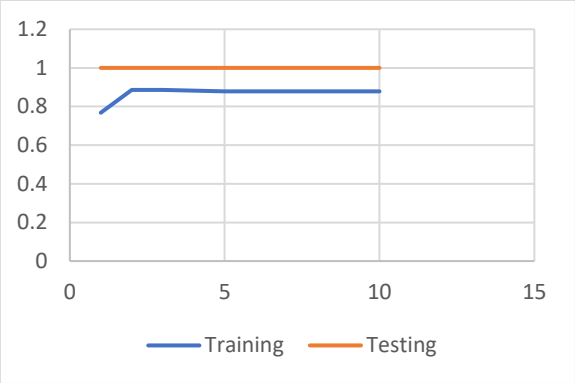
10 Trees



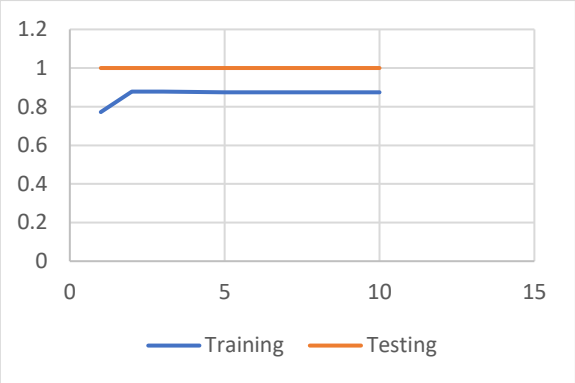
20 Trees



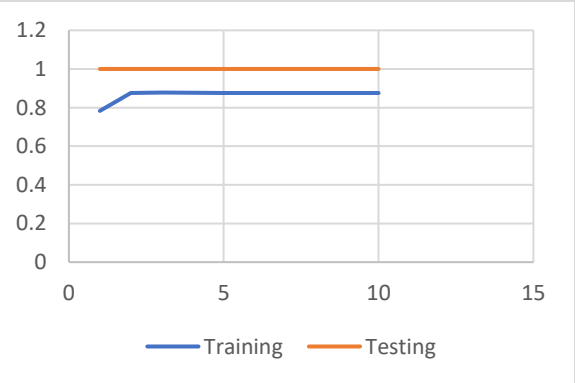
40 Trees



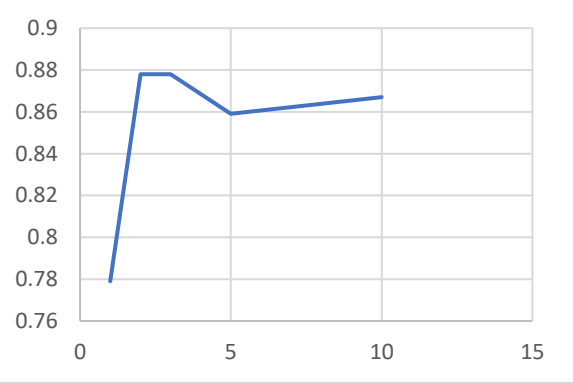
60 Trees



80 Trees



100 Trees



SVM Classification learner (weka.classifiers.functions.supportVector). You will run the SVM classifier on the training data to answer the following questions.

(a) Using a linear kernel (-t 0 option), train the SVM on the training data for different values of C parameter(-c option): 10^−4 , 10^−3 , 10^−2 , 10^−1 , 10^0 , 10^1 , 10^2 , 10^3 , 10^4 . Compute the training accuracy, and testing accuracy for the SVM obtained with different values of the C parameter. Plot the training accuracy and testing accuracy as a function of C (C value on x-axis and Accuracy on y-axis) – *one curve each for training, validation, and testing data*. List your observations.

(b) Repeat the experiment (a) with polynomial kernel (-t 1 -d option) of degree 2, 3, and 4. Compare the training and testing accuracies for different kernels (linear, polynomial kernel of degree 2, polynomial kernel of degree 3, and polynomial kernel of degree 4). List your observations.

Voting

a. Linear Kernel

Degree	C Parameter	Training Accuracy	Testing Accuracy
0	0.0001	0.621	0.590
0	0.001	0.621	0.590
0	0.01	0.949	0.890
0	0.1	0.967	0.920
0	1	0.976	0.930
0	10	0.979	0.920
0	100	0.976	0.910
0	1000	0.976	0.910
0	10000	0.976	0.910

b. Poly Degree

Degree	C Parameter	Training Accuracy	Testing Accuracy
2	0.0001	0.621	0.590
2	0.001	0.946	0.900
2	0.01	0.970	0.930
2	0.1	0.982	0.950
2	1	0.997	0.920
2	10	0.997	0.930
2	100	0.997	0.930
2	1000	0.997	0.930
2	10000	0.997	0.930

Degree	C Parameter	Training Accuracy	Testing Accuracy
3	0.0001	0.940	0.890
3	0.001	0.976	0.920
3	0.01	0.988	0.950
3	0.1	0.997	0.930
3	1	0.997	0.930
3	10	0.997	0.930
3	100	0.997	0.930
3	1000	0.997	0.930
3	10000	0.997	0.930

Degree	C Parameter	Training Accuracy	Testing Accuracy
4	0.0001	0.979	0.930
4	0.001	0.988	0.950
4	0.01	0.997	0.930
4	0.1	0.997	0.930
4	1	0.997	0.930
4	10	0.997	0.930
4	100	0.997	0.930
4	1000	0.997	0.930
4	10000	0.997	0.930

Ionosphere

a. Linear kernel

Degree	C Parameter	Training Accuracy	Testing Accuracy
0	0.0001	0.521	1.000
0	0.001	0.521	1.000
0	0.01	0.521	1.000
0	0.1	0.521	1.000
0	1	0.521	1.000
0	10	0.521	1.000
0	100	0.521	1.000
0	1000	0.521	1.000
0	10000	0.521	1.000

b. Poly Degree

Degree	C Parameter	Training Accuracy	Testing Accuracy
2	0.0001	0.521	1.000
2	0.001	0.802	1.000
2	0.01	0.871	1.000
2	0.1	0.867	1.000
2	1	0.878	1.000
2	10	0.856	1.000
2	100	0.821	1.000
2	1000	0.821	1.000
2	10000	0.821	1.000

Degree	C Parameter	Training Accuracy	Testing Accuracy
3	0.0001	0.844	1.000
3	0.001	0.859	1.000
3	0.01	0.867	1.000
3	0.1	0.859	1.000
3	1	0.829	1.000
3	10	0.821	1.000
3	100	0.821	1.000
3	1000	0.821	1.000
3	10000	0.821	1.000

Degree	C Parameter	Training Accuracy	Testing Accuracy
4	0.0001	0.871	1.000
4	0.001	0.875	1.000
4	0.01	0.863	1.000
4	0.1	0.825	1.000
4	1	0.825	1.000
4	10	0.825	1.000
4	100	0.825	1.000
4	1000	0.825	1.000
4	10000	0.825	1.000